Uncertainty and Conflict: A Behavioural Approach to the Aggregation of Expert Opinions

Matthias C. M. Troffaes
SYSTeMS Research Group
Ghent University
Matthias.Troffaes@UGent.be

Abstract

A second-order imprecise probability model is proposed to generalise the conjunction rule in case of expert conflict. The essential idea underlying the model is a notion of behavioural trust. A computationally feasible algorithm for calculating the first-order aggregate is constructed.

1 Introduction

When modelling a system, one must often rely on expert information. From the modeller's perspective, one usually wants to aggregate all expert opinions into a single representative model—a "summary" of all the expert information—which must then serve as a basis for various kinds of inferences about the system, such as decision making, estimation, hypothesis testing, *etc.* The fundamental idea underlying this approach is that aggregating more expert opinions eventually leads to a more reliable, and hopefully, also a more informative representative model. There is however no agreement on how expert opinions should be aggregated. Actually, there is not even a clear agreement on how expert opinions themselves should be represented.

Recently, the use of imprecise probabilities in representing, manipulating and aggregating expert information has received an increasing amount of attention in the literature (see for instance [15, 16, 19, 17, 10, 4, 11, 6, 13, 14, 5] and many references therein). Some of the main reasons for the increasing popularity of imprecise probabilities in modelling and aggregating expert information are that they (i) allow for a more reliable representation of expert information (essentially, they do not force the expert to pinpoint a single probability measure in order to represent his knowledge), and (ii) provide a natural setting for modelling conflicting opinions, using imprecision as a means of expressing disagreement amongst different opinions.

The main goal of this paper is to provide a transparent, systematic and computationally feasible way for reconciling conflicting expert opinions. I investigate how an imprecise second-order hierarchical model can be constructed, how this model can be given an operational meaning, and last but not least, I derive an efficient algorithm for calculating a first order aggregate from the second-order model, but not for all cases. I will show that for some, even quite simple, second-order models this is not necessarily possible. Throughout this investigation I shall try to *use behavioural arguments only*, in particular, avoiding sure loss, coherence and natural extension, which are the fundamental concepts of the behavioural theory of imprecise probabilities [18]. The proposed method generalises a second-order hierarchical model described in [3, 4] and is mathematically closely related to results presented in [13].

The paper is organised as follows. Section 2 introduces the basic concepts of the behavioural theory of imprecise probabilities under the form of lower previsions, and their relation to other well-known uncertainty models. In Section 3 I briefly review the problem of aggregating expert opinions, and touch on the controversy surrounding it. Section 4 explains the conjunction rule. A second-order imprecise probability model is proposed and discussed in Section 5. In Section 6 the main results and an example are presented, and I end with a discussion in Section 7.

2 Lower previsions

In this paper, lower previsions are taken as the fundamental imprecise probability model. Its behavioural interpretation turns out to be very convenient in describing the second-order model later on.

Let us consider a subject (which can be an expert, or a modeller) who is uncertain about something, say, the outcome of some experiment. If the set of possible outcomes is $\mathcal{A}$, then a *gamble* X is a bounded mapping from $\mathcal{A}$ to $\mathbb{R}$, and it is interpreted as an uncertain reward: if a turns out to be the true outcome of the experiment then the subject receives the amount $X(a)$, expressed in units of some linear utility. The set of all gambles on $\mathcal{A}$ is denoted by $\mathcal{L}(\mathcal{A})$.

The information the subject has about the outcome of the experiment will lead him to accept or reject transactions whose reward depends on this outcome, and we can formulate a model for his uncertainty by looking at a specific type of transaction: the buying of gambles. The subject's *lower prevision* (or supremum acceptable buying price) $\underline{P}(X)$ for a gamble X is the highest price s such that he is disposed to buy the gamble X for any price strictly lower than s. If the subject assesses a supremum acceptable buying price for every gamble X in a subset $\mathcal{K}$ of $\mathcal{L}(\mathcal{A})$, the resulting mapping $\underline{P} : \mathcal{K} \to \mathbb{R}$ is called a *lower prevision.*

Examples of lower previsions are:

(i) If "a belongs to the set $A \subseteq \mathcal{A}$" then $\underline{P}_A(X) = \inf_{a \in A} X(a)$: the lowest possible reward given that $a \in A$. We call $\underline{P}_A$ the *vacuous lower prevision* relative to A.

(ii) If "a has probability density ϕ" we should pay $P(X) = \int_{\mathcal{A}} X(a)\phi(a)\mathrm{d}a$, the expectation w.r.t. ϕ [7, 8]. This is called the *linear prevision* induced by the density ϕ.

(iii) If "a has a probability density that belongs to the set Φ" we pay at most $\underline{P}(X) = \inf_{\phi\in\Phi} \int_{\mathcal{A}} X(a)\phi(a)\mathrm{d}a$.

These examples indicate that lower previsions are uncertainty *representations* that are expressive enough to capture propositional logic (example (i)), Bayesian probability theory (example (ii)), and credal sets (example (iii)) (credal sets are convex sets of probability measures). Actually, they also generalise belief functions, possibility and necessity measures, Choquet capacities, risk measures, and many others (for more details see [18]). Lower previsions may therefore be seen as a unifying imprecise probability framework.

$\overline{P}$ will denote the conjugate *upper prevision* of $\underline{P}$. It is defined by $\overline{P}(X) = -\underline{P}(-X)$ for every $X \in -\mathcal{K}$. $\overline{P}(X)$ represents the subject's infimum acceptable selling price for the gamble X. The difference $\overline{P}(X) - \underline{P}(X)$ is a measure for the amount of imprecision in the subject's behavioural dispositions towards X.

We now introduce a method of *inference*, associated with lower previsions, that also generalises the inference methods of, for instance, classical propositional logic and Bayesian probability theory.

2.1 Inference

Through a procedure called *natural extension*, we are able to derive from the assessments embodied in $\underline{P}$, a supremum buying price $\underline{E}(X)$ for each gamble X in $\mathcal{L}(\mathcal{A})$; $\underline{E}$ is the smallest (and therefore most conservative) lower prevision that satisfies, for any gambles X and Y

- $\underline{E}(X) \geq \inf[X]$ (accepting sure gain)
- $\underline{E}(\lambda X) = \lambda\underline{E}(X)$ whenever $\lambda > 0$ (scale independence)
- $\underline{E}(X + Y) \geq \underline{E}(X) + \underline{E}(Y)$ (super-additivity)
- $\underline{E}(X) \geq \underline{P}(X)$ (compatibility)

If $\underline{E}$ exists, $\underline{P}$ is said to *avoid sure loss*. It can be easily shown that a lower prevision avoids sure loss if and only if $\sup_{a\in\mathcal{A}} \left[\sum_{i=1}^{n} [X_i(a) - \underline{P}(X_i)]\right] \geq 0$ for any $n \in \mathbb{N}$ and any $X_1, \ldots, X_n \in \mathcal{K}$; that is, if and only if there is no combination of transactions—buying gambles for their supremum buying price—that leads to a sure loss.

The natural extension $\underline{E}(X)$ can be easily calculated: assuming $\mathcal{K}$ to be finite, it is equal to the supremum achieved by the free variable α subject to

$$X(a) - \alpha \geq \textstyle\sum_{Y\in\mathcal{K}} \lambda_Y \big(Y(a) - \underline{P}(Y)\big)$$

for each $a \in \mathcal{A}$, with variables $\lambda_Y \geq 0$ for each $Y \in \mathcal{K}$—if also $\mathcal{A}$ is finite,[1] this is a linear program. If the supremum is $\alpha = +\infty$, then the natural extension

[1] It happens very often in practice that both $\mathcal{K}$ and $\mathcal{A}$ are finite.

does not exist, and hence, $\underline{P}$ incurs sure loss; this identifies a conflict in the assessments. If $\underline{E}$ and $\underline{P}$ coincide on $\mathcal{K}$, then $\underline{P}$ is called *coherent*.

3 Aggregation: a short review

I now give a short non-exhaustive review of different ways to tackle the problem of aggregating expert opinions. Basically, there are two ways to approach the problem: *axiomatic* (also called normative), and *ad hoc*. No rule is ever purely axiomatic, or purely *ad hoc*. Many rules can be given an axiomatic as well as an *ad hoc* explanation (such as the conjunction rule and the unanimity rule described below).

Axiomatic approaches aim at deriving a preferably unique rule of aggregation from axioms or properties that this rule should satisfy. Typical axioms are requirements of commutativity of the rule with respect to some other action, such as updating (external Bayesianity), marginalisation, permutation of experts (symmetry) *etc.*. They can also refer to some other property of the rule, such as unanimity-preservation (if all experts agree, then the aggregate should also agree with all experts), invariance with respect to non-informative expert opinions, independence preservation, *etc.*

Especially among Bayesians (see [9] for an excellent overview, and references therein), where expert opinions and the aggregate are to be represented by probability measures, there still is a lot of controversy about these axioms. Indeed, imposing even only a few axioms easily leads to contradictions or undesirable aggregation rules such as so-called dictatorship rules. What counts is how the rule will eventually be used. From this perspective, it is not always clear what axioms should be imposed.

In imprecise probability theory, the axiomatic approach is somewhat less problematic (see [15] for instance). Still, it is not clear how to define a unique aggregation rule under this uncertainty model. The *conjunction rule* is defined as the smallest (and therefore most conservative) coherent lower prevision that dominates each of the experts' lower previsions. Conjunction aims at gaining as much information as possible from each of the experts: the aggregate is at least as informative as each of the experts' lower previsions, and it can only become more informative as more experts enter the scene. The conjunction however does not always exists, in particular when different experts make conflicting statements. On the other hand, the *unanimity rule*, defined as the lower envelope of the experts' lower previsions, is guaranteed to exist. It aims at reconciling the experts' assessments. As a result however, it may lead to extremely imprecise results: the aggregate will be at least as imprecise as the most imprecise expert, and its imprecision can only increase as more experts enter the scene. Unanimity certainly leads to a very reliable aggregate. However, it fails completely to produce also a more informative aggregate as more expert assessments become available.

One imprecise probability aggregation rule could consist of using the conjunction rule if the conjunction exists, and the unanimity rule if the conjunction does not exist. The problem of this rule is that it is far from stable: a small variation of an expert's lower prevision may yield huge differences in the aggregate lower prevision.

Ad hoc approaches are not as much concerned with axioms: one simply proposes or derives a mathematical formula, together with some form of justification. (Afterwards of course, it is usually investigated which of the axioms it satisfies. This usually provides the *ad hoc* rule with an additional source for motivation or criticism.) They generally divide into three sub-categories: hierarchical models, weighting schemes, and consensus methods. Consensus methods are based on expert interaction: before an aggregate is constructed, the experts are allowed to interact with each other (see [11] for an excellent discussion of a consensus method using imprecise probabilities). Weighting rules, the linear opinion pool in Bayesian aggregation being maybe the most prevailing example, try to take each expert's expertise into account (a feature lacking most of the purely axiomatic approaches). The same holds for hierarchical models, and in fact, hierarchical models may be seen as one attempt to motivate, and generalise, some of the existing weighting schemes (many weighting rules are however not instances of hierarchical models).

Using probability measures, the most reasonable approach seems to be linear pooling; taking a convex combination of expert probability measures. It is very easy to implement, and gives quite good results in practice. Subject of debate is of course how one should assign the weights.

Concluding, besides theoretical and practical problems associated with each of these methods separately, any method using single probability distributions for both the experts and the aggregate fails to model conflict among experts, and forces experts to pinpoint a single probability, even for those events of which he does not have much expertise. Imprecise probabilities address both these problems, because they allow for experts to assess their expertise using a convex set of probability measures (also called credal sets), a lower prevision, a set of desirable gambles, an ordering on gambles, a possibility measure, *etc.*—rather than forcing them to choose a single probability measure. Consequently, it is also easier to avoid conflict when combining imprecise probabilities because, roughly speaking, experts are not forced to give precise probabilities on events of which they have only little knowledge—they can simply say they don't know. And should there be conflict anyway, imprecision can be used to reflect it (for instance, using the unanimity rule). These characteristics are the main motivation for introducing the second-order imprecise probability model in Section 5.

4 Prelude: the conjunction rule

Suppose there are N (male) subjects, called *experts.* Suppose that their assessments about the value that a parameter ω assumes in a *finite* set of possible values Ω are expressed through coherent lower previsions $\underline{P}_k$ on some *finite* subset $\mathcal{K}_k$ of $\mathcal{L}(\Omega)$, for $k = 1, \dots, N$. The natural extension of each $\underline{P}_k$ will be denoted by $\underline{E}_k$. How can the lower previsions $(\underline{P}_k)_{k=1}^N$ be combined into an aggregate, a single coherent lower prevision defined on the set of gambles $\mathcal{L}(\Omega)$?

Consider therefore a new (female) subject, called the *modeller*. She wishes to aggregate the expert assessments to a single coherent lower prevision $\underline{P}_M$ defined on $\mathcal{L}(\Omega)$. Let us first introduce a notion of behavioural trust.

Definition 1. *Let α and β be two subjects. Assume that each of the subjects models his/her knowledge about $\omega \in \Omega$ through a coherent lower prevision $\underline{P}_\alpha$ resp. $\underline{P}_\beta$ on $\mathcal{K}_\alpha$ resp. $\mathcal{K}_\beta$. Let $\underline{E}_\alpha$ resp. $\underline{E}_\beta$ denote their natural extension. The following conditions are equivalent; if any (hence all) of them are satisfied, we say that α* trusts *β.*

(A) α is willing to accept every decision β makes concerning buying gambles on Ω, that is, for each gamble $X \in \mathcal{L}(\Omega)$, α is willing to accept β's price $s < \underline{E}_\beta(X)$ for buying X as his/her price for buying X.

(B) $\underline{E}_\alpha$ point-wise dominates $\underline{E}_\beta$ on $\mathcal{L}(\Omega)$.

The point of the first part of the definition is that *any behavioural theory of uncertainty inherently has a notion of trust in a multi-agent environment and hence,* as I will show now, *also notions of conjunction and consistency*, which can be derived from behavioural trust in a straightforward way.

Definition 2. *The* conjunction *of $(\underline{P}_k)_{k=1}^N$ is defined as the smallest, and hence most conservative, lower prevision $\underline{P}_M$ on $\mathcal{L}(\Omega)$ the modeller can have such that she still trusts each of the experts. If conjunction exists, then the experts are said to be* consistent*, otherwise they are said to be* conflicting.

By Definition 1, the conjunction is simply the (point-wise) smallest coherent lower prevision that dominates all the experts' natural extensions $\underline{E}_k$. The conjunction of $(\underline{P}_k)_{k=1}^N$ be denoted by $\sqcap_{k=1}^N \underline{P}_k$; the conjunction of two consistent coherent lower previsions $\underline{P}_1$ and $\underline{P}_2$ is also denoted by $\underline{P}_1 \sqcap \underline{P}_2$. It is easy to show that $\sqcap$ is an associative and commutative operator on coherent lower previsions (but the result is only defined in case of consistency). Conjunction can be calculated through linear programming in a similar way as natural extension.

Proposition 1. *Consider the maximum α^* achieved by the free variable α subject to the linear constraints*

$$X(a) - \alpha \geq \sum_{k=1}^{N} \sum_{Y_k \in \mathcal{K}_k} \lambda_{Y_k} \big(Y_k(a) - \underline{P}_k(Y_k) \big)$$

for each $a \in \mathcal{A}$, with variables $\lambda_{Y_k} \geq 0$ for each $Y_k \in \mathcal{K}_k$. If α^* is finite then the experts' assessments $(\underline{P}_k)_{k=1}^N$ are consistent and $\sqcap_{k=1}^N \underline{P}_k = \alpha^*$. If $\alpha^* = +\infty$, then the conjunction does not exist: in such a case the assessments $(\underline{P}_k)_{k=1}^N$ are conflicting.

If the assessments $(\underline{P}_k)_{k=1}^N$ are conflicting—if no conjunction exists—then there is no coherent way to accept *every* decision of *every* expert, since the modeller incurs a sure loss if she would do so. It is easily established that in case of inconsistency there are gambles $X_k \in \mathcal{K}_k$ such that (compare with avoiding sure loss)

$$\sup_{\omega \in \Omega} \left[\sum_{k=1}^{N} [X_k(\omega) - \underline{P}_k(X_k)] \right] < 0, \tag{1}$$

i.e., the combination of the transactions in which the gambles X_k are bought for a price $\underline{P}_k(X_k)$ leads to a loss, whatever the actual value of the parameter a. Blindly accepting decisions of all the experts $(\underline{P}_k)_{k=1}^N$ is clearly unacceptable in case of inconsistency. The modeller is therefore certain that some of the experts' assessments $(\underline{P}_k)_{k=1}^N$ cannot be trusted, but she does not necessarily know which ones.

One solution in case of conflict is to use the unanimity rule. This consists in choosing the modeller's lower prevision, the aggregate, such that each of the experts trust the modeller. This means that each of the experts agrees with the modeller's behavioural dispositions (hence the name of the rule). But as we have already noted before, the resulting aggregate may be too imprecise to be useful.

It may however happen that the modeller may have actual information about which of the experts are to be trusted more than others. In the next section I therefore propose a second-order hierarchical imprecise probability model that aims at modelling such knowledge. Its interpretation is based on the notion of behavioural trust.

5 A second-order imprecise probability model

The modeller wishes to recover information regarding ω using the information revealed by the experts, taking into account that some experts are more trustworthy than others. I describe how behavioural trust can be used to aggregate information revealed by experts.

The modeller first assumes the existence of a so-called *true coherent lower prevision* $\underline{P}_T$ on $\mathcal{L}(\Omega)$, but she is not sure about what it is. $\underline{P}_T$ could refer to the behaviour of a hypothetical "representative" expert, an operational procedure designed to measure uncertainty such as an imprecise Dirichlet (or other) model updated through a contingency table, or even a real system that behaves just like an expert. The modeller is interested in what the hypothetical expert knows about ω, or what the result of the operational procedure will be about ω, or how the system behaves with respect to ω, but, she is only able to infer information about ω through $(\underline{P}_k)_{k=1}^N$. She cannot talk to the hypothetical representative expert, cannot perform operational procedure, has no access to the system of interest: it may be too expensive, or she might not have the necessary means. Her uncertainty thus regards the random variable $\underline{P}_T$ which we assume to take all values in the set $\underline{\mathcal{P}}(\Omega)$ of coherent lower previsions on $\mathcal{L}(\Omega)$. Her possibility space $\underline{\mathcal{P}}(\Omega)$ is also called the second-order possibility space.[2]

Often, even in imprecise probability theory, the second-order possibility space is restricted to the set of all linear previsions. It is well-known that this restriction may lead to different results: precision-imprecision equivalence does not always hold [4]. My main motivation for *not* restricting to linear previsions is that we should not expect experts to be able to pinpoint a single probability measure. We want experts to be honest about their information, so if there really is uncertainty, we sure want them to be able to tell us. This should hold as well for the "real" experts as for the hypothetical representative expert, operational procedure, or real system.

5.1 Trust and tsurt

In terms of events on the modeller's second-order possibility space, she trusts an expert, with lower prevision $\underline{P}_k$, whenever the event $\underline{\mathcal{M}}(\underline{E}_k)$ is true (remember that $\underline{E}_k$ is the natural extension of $\underline{P}_k$), with

$$\underline{\mathcal{M}}(\underline{E}_k) = \{\underline{P}_T \in \underline{\mathcal{P}}(\Omega) \colon (\forall X \in \mathcal{L}(\Omega))(\underline{E}_k(X) \leq \underline{P}_T(X))\}.$$

Indeed, this event is true exactly when the true lower prevision point-wise dominates $\underline{E}_k$; this means that the true behavioural dispositions, implied by $\underline{P}_T$, include at least the behavioural dispositions implied by $\underline{E}_k$: no harm is done to the modeller by being guided by the decisions the expert makes.

Dually, the modeller might think that behavioural dispositions implied by the expert's natural extension $\underline{E}_k$, include at least the behavioural dispositions implied by $\underline{P}_T$. The modeller judges the expert's assessment to be too precise (for instance, he might be a Bayesian), but not necessarily contradicting $\underline{P}_T$. Let us say in such a case that the modeller is trusted by, or *tsurts* the expert.[3] In terms of events on the modeller's second-order possibility space, the modeller tsurts an expert whenever $\underline{\mathcal{N}}(\underline{E}_k)$ is true, with

$$\underline{\mathcal{N}}(\underline{E}_k) = \{\underline{P}_T \in \underline{\mathcal{P}}(\Omega) \colon (\forall X \in \mathcal{L}(\Omega))(\underline{P}_T(X) \leq \underline{E}_k(X))\}.$$

(note that $\underline{\mathcal{N}}(\underline{E}_k) \neq \complement\underline{\mathcal{M}}(\underline{E}_k)$).[4]

The modeller assesses a supremum betting rate t_k^ℓ for the event that she can trust expert k, i.e., for the event $\underline{\mathcal{M}}(\underline{E}_k)$ and a supremum betting rate $1 - t_k^u$ for the event that she cannot trust expert k, i.e., for the event $\complement\underline{\mathcal{M}}(\underline{E}_k)$. (The interval $[t_k^\ell, t_k^u]$ can be interpreted as a probability interval for the trust in expert k.)

Similarly, the modeller assesses a supremum betting rate v_k^ℓ for the event that expert k trusts her, i.e., for the event $\underline{\mathcal{N}}(\underline{E}_k)$ and a supremum betting rate $1 - v_k^u$ for the event that expert k does not trust her, i.e., for the event $\complement\underline{\mathcal{N}}(\underline{E}_k)$.

[2]The first-order possibility space is Ω, and $(\underline{P}_k)_{k=1}^N$ and $\underline{P}_T$ are called first-order models.

[3]Out of convenience we define the verb *to tsurt [somebody]*: being trusted [by somebody], and the noun *tsurt*: dispositions you have when you tsurt someone.

[4]The symbol $\complement$ denotes the complement of a set.

(The interval $[v_k^\ell, v_k^u]$ can be interpreted as a probability interval for the tsurt of expert k.)

Define *lower & upper trust functions* and *lower & upper tsurt functions* as follows:

$$\underline{\mathfrak{t}}\colon \underline{E}_k \mapsto t_k^\ell, \quad \overline{\mathfrak{t}}\colon \underline{E}_k \mapsto t_k^u, \quad \underline{\mathfrak{v}}\colon \underline{E}_k \mapsto v_k^\ell, \quad \overline{\mathfrak{v}}\colon \underline{E}_k \mapsto v_k^u.$$

Without loss of generality we may assume that $\underline{\mathfrak{t}}$, $\overline{\mathfrak{t}}$, $\underline{\mathfrak{v}}$ and $\overline{\mathfrak{v}}$ are defined on a common domain $\{\underline{E}_k \colon k \in \{1, \ldots, n\}\}$: the modeller can always choose betting rate 0 for events of which she is completely ignorant. In terms of probability intervals this means that lower probabilities can be chosen 0 and the upper probabilities can be chosen 1 whenever they are unknown.

Obviously it should hold that $t_k^\ell \leq t_k^u$ and $v_k^\ell \leq v_k^u$. Extreme choices are $t_k^\ell = 1$, this corresponds to complete trust, and $v_k^\ell = 1$, which corresponds to complete tsurt.

We defined trust and tsurt functions $\underline{\mathfrak{t}}$, $\overline{\mathfrak{t}}$, $\underline{\mathfrak{v}}$ and $\overline{\mathfrak{v}}$ as specifications of a supremum betting rate or lower probability $\underline{Q}$ on particular events in the second order possibility space $\underline{\mathcal{P}}(\Omega)$. In terms of this second order possibility space, the modeller has specified the following lower prevision:

$$\begin{aligned} \underline{Q}(\underline{\mathcal{M}}(\underline{E}_k)) &= \underline{\mathfrak{t}}(\underline{E}_k), & \underline{Q}(\complement\underline{\mathcal{M}}(\underline{E}_k)) &= 1 - \overline{\mathfrak{t}}(\underline{E}_k), \\ \underline{Q}(\underline{\mathcal{N}}(\underline{E}_k)) &= \underline{\mathfrak{v}}(\underline{E}_k), & \underline{Q}(\complement\underline{\mathcal{N}}(\underline{E}_k)) &= 1 - \overline{\mathfrak{v}}(\underline{E}_k). \end{aligned}$$

It is now convenient to define the mapping $\delta\colon \underline{\mathcal{P}}(\Omega) \times \underline{\mathcal{P}}(\Omega) \to \{0, 1\}$ by

$$\delta_{\underline{P}_1}^{\underline{P}_2} = \begin{cases} 1, & \text{if for each } X \in \mathcal{L}(\Omega)\colon \underline{P}_1(X) \leq \underline{P}_2(X), \\ 0, & \text{otherwise.} \end{cases}$$

Observe that $\underline{\mathcal{M}}(\underline{E}_k)(\underline{P}) = \delta_{\underline{E}_k}^{\underline{P}}$ and $\underline{\mathcal{N}}(\underline{E}_k)(\underline{P}) = \delta_{\underline{P}}^{\underline{E}_k}$.[5]

5.2 A first-order aggregate through natural extension

If $\underline{Q}$ avoids sure loss, that is, if

$$\sup_{\underline{P} \in \underline{\mathcal{P}}(\Omega)} \left\{ \sum_{k=1}^{N} \kappa_k \Big(\delta_{\underline{E}_k}^{\underline{P}} - \underline{\mathfrak{t}}(\underline{E}_k) \Big) + \lambda_k \Big(\overline{\mathfrak{t}}(\underline{E}_k) - \delta_{\underline{E}_k}^{\underline{P}} \Big) \right. \\ \left. + \mu_k \Big(\delta_{\underline{P}}^{\underline{E}_k} - \underline{\mathfrak{v}}(\underline{E}_k) \Big) + \nu_k \Big(\overline{\mathfrak{v}}(\underline{E}_k) - \delta_{\underline{P}}^{\underline{E}_k} \Big) \right\} \geq 0,$$

for every κ_k, λ_k, μ_k and $\nu_k \geq 0$, then the natural extension $\underline{E}$ of $\underline{Q}$ exists; it is a coherent lower prevision on $\mathcal{L}(\underline{\mathcal{P}}(\Omega))$. In such a case we say that there is

[5] We identify a set with its indicator function, for instance, a subset M of $\underline{\mathcal{P}}(\Omega)$ is identified with the mapping $M(\underline{P}) = \begin{cases} 0, & \text{if } \underline{P} \notin M, \\ 1, & \text{otherwise.} \end{cases}$

second-order consistency. If $\underline{Q}$ does not avoids sure loss, then we say that there is *second-order conflict.* The natural extension, if it exists, is given by

$$\underline{E}(Z) = \sup \Bigg\{ \alpha \in \mathbb{R} \colon (\exists \kappa_k, \lambda_k, \mu_k, \nu_k \geq 0)(\forall \underline{P} \in \underline{\mathcal{P}}(\Omega))$$
$$Z(\underline{P}) - \alpha \geq \sum_{k=1}^{N} \kappa_k \left(\delta^{\underline{P}}_{\underline{E}_k} - \underline{\mathfrak{t}}(\underline{E}_k) \right) + \lambda_k \left(\overline{\mathfrak{t}}(\underline{E}_k) - \delta^{\underline{P}}_{\underline{E}_k} \right)$$
$$+ \mu_k \left(\delta^{\underline{E}_k}_{\underline{P}} - \underline{\mathfrak{v}}(\underline{E}_k) \right) + \nu_k \left(\overline{\mathfrak{v}}(\underline{E}_k) - \delta^{\underline{E}_k}_{\underline{P}} \right) \Bigg\}$$

for any (second-order) gamble $Z \in \mathcal{L}(\underline{\mathcal{P}}(\Omega))$.

From the natural extension $\underline{E}$, we can, theoretically, deduce lower and upper trust, and lower and upper tsurt of any coherent lower prevision $\underline{P}$:

$$\underline{\mathfrak{t}}^e(\underline{P}) = \underline{E}(\underline{\mathcal{M}}(\underline{P})), \qquad \overline{\mathfrak{t}}^e(\underline{P}) = 1 - \underline{E}(\complement \underline{\mathcal{M}}(\underline{P})),$$
$$\underline{\mathfrak{v}}^e(\underline{P}) = \underline{E}(\underline{\mathcal{N}}(\underline{P})), \qquad \overline{\mathfrak{v}}^e(\underline{P}) = 1 - \underline{E}(\complement \underline{\mathcal{N}}(\underline{P})).$$

These natural extensions will agree with the original assessments if $\underline{Q}$ is coherent (the formula expressing this is rather lengthy). We may also infer supremum buying prices and infimum selling prices for the supremum buying price and infimum selling price of a gamble X with respect to true model:

$$\underline{E}^{\ell}(X) = \underline{E}(X_*), \qquad \underline{E}^{u}(X) = \underline{E}(X^*),$$
$$\overline{E}^{\ell}(X) = \overline{E}(X_*), \qquad \overline{E}^{u}(X) = \overline{E}(X^*),$$

where X_* is the lower and X^* is the upper evaluation map corresponding to X, defined by

$$X_* \colon \underline{\mathcal{P}}(\Omega) \to \mathbb{R}; \underline{P} \mapsto \underline{P}(X), \quad X^* \colon \underline{\mathcal{P}}(\Omega) \to \mathbb{R}; \underline{P} \mapsto \overline{P}(X).$$

$\underline{E}^{\ell}$, $\underline{E}^{u}$, $\overline{E}^{\ell}$ and $\overline{E}^{u}$ are all first-order models, but which one should we choose as first-order aggregate? $\underline{E}^{u}$ and $\overline{E}^{\ell}$ are not necessarily coherent lower previsions, but $\underline{E}^{\ell}$ and $\underline{E}^{u}$ are. The next proposition shows that it makes sense to take $\underline{E}^{\ell}$ as the first order aggregate.

Proposition 2. *$\underline{E}^{\ell}$ is a coherent lower prevision and $\underline{E}^{u}$ is a coherent upper prevision. Moreover, for every gamble X it holds that $\underline{E}^{\ell}(X) = -\overline{E}^{u}(-X)$.*

6 Main result

The following theorem establishes that whenever $\overline{\mathfrak{v}}$ is 1, we can *efficiently* calculate $\underline{E}^{\ell}(X)$ for every gamble X, and $\underline{\mathfrak{t}}^e(\underline{Q})$ and $\overline{\mathfrak{v}}^e(\underline{Q})$ for every coherent lower prevision $\underline{Q}$, by taking for Z either X_*, $\underline{\mathcal{M}}(\underline{Q})$ or $1 - \underline{\mathcal{N}}(\underline{Q})$.

Theorem 1. *Suppose that $Z \in \mathcal{L}(\underline{\mathcal{P}}(\Omega))$ is monotonically increasing and $\overline{\mathfrak{v}}(\underline{E}_k) = 1$ for every $k \in \{1, \dots, N\}$. Then*

$$\begin{aligned}\underline{E}(Z) = \sup\Bigg\{\alpha \in \mathbb{R}\colon &\\ &(\exists \kappa_k, \lambda_k, \mu_k \geq 0)(\forall K \subseteq \{1, \dots, N\})(\underline{R} = \sqcap_{k \in K} \underline{P}_k \text{ coherent})\\ &Z(\underline{R}) - \alpha \geq \sum_{k=1}^{N} \kappa_k \Big(\delta^{\underline{R}}_{\underline{E}_k} - \underline{\mathfrak{t}}(\underline{E}_k)\Big) + \lambda_k \Big(\overline{\mathfrak{t}}(\underline{E}_k) - \delta^{\underline{R}}_{\underline{E}_k}\Big)\\ &\qquad\qquad + \mu_k \Big(\delta^{\underline{E}_k}_{\underline{R}} - \underline{\mathfrak{v}}(\underline{E}_k)\Big)\Bigg\}\end{aligned}$$

(where $\underline{R} = \sqcap_{k \in K} \underline{P}_k$ is defined as the vacuous lower prevision $\inf_{\omega \in \Omega}$ *if $K = \emptyset$.)*

Proof. If $\overline{\mathfrak{v}}(\underline{E}_k) = 1$ for all k then the supremum will be achieved for $\nu_k = 0$, and hence, we may omit these terms.

Next, we show that for any $\underline{P} \in \underline{\mathcal{P}}(\Omega)$ we can find a K such that,

$$Z(\underline{R}) \leq Z(\underline{P}), \qquad \delta^{\underline{R}}_{\underline{E}_k} = \delta^{\underline{P}}_{\underline{E}_k}, \qquad \delta^{\underline{E}_k}_{\underline{R}} \geq \delta^{\underline{E}_k}_{\underline{P}},$$

for all $k \in \{1, \dots, N\}$, with $\underline{R} = \sqcap_{k \in K} \underline{P}_k$. In such a case the inequality for $\underline{P}$ is implied by the inequality for $\underline{R}$, and hence, we may 'replace' $\underline{P}$ by $\underline{R}$ in the inequality.

Choose $K = \{k\colon \underline{E}_k \leq \underline{P}\}$. Observe that $\underline{R} = \sqcap_{k \in K} \underline{P}_k$ is a coherent lower prevision (if $K = \emptyset$ then $\underline{R}$ is the vacuous lower prevision $\inf_{\omega \in \Omega}$).

Also observe that $\underline{R} \leq \underline{P}$, and hence, it immediately follows that $Z(\underline{R}) \leq Z(\underline{P})$ since Z is monotone, and $\delta^{\underline{R}}_{\underline{E}_k} \leq \delta^{\underline{P}}_{\underline{E}_k}$ and $\delta^{\underline{E}_k}_{\underline{R}} \geq \delta^{\underline{E}_k}_{\underline{P}}$ for every $k \in \{1, \dots, N\}$, since

$$\begin{aligned}\underline{R} \geq \underline{E}_k &\implies \underline{P} \geq \underline{E}_k,\\ \underline{E}_k \geq \underline{P} &\implies \underline{E}_k \geq \underline{R}.\end{aligned}$$

We are left to show that

$$\underline{P} \geq \underline{E}_k \implies \underline{R} \geq \underline{E}_k,$$

which would establish $\delta^{\underline{R}}_{\underline{E}_k} = \delta^{\underline{P}}_{\underline{E}_k}$. Indeed, suppose that $\underline{P} \geq \underline{E}_k$. This means that $k \in K$. Since $\underline{R} \geq \underline{E}_\ell$ for all $\ell \in K$ by definition of $\underline{R}$, we find that $\underline{R} \geq \underline{E}_k$. □

We must require that $\overline{\mathfrak{v}}$ is 1 because in general it is impossible to establish that for every $\underline{P}$ there is a $\underline{R}$, chosen from a finite set constructed from $\{\underline{P}_k, k \in \{1, \dots, N\}\}$, such that

$$Z(\underline{R}) \leq Z(\underline{P}), \qquad \delta^{\underline{R}}_{\underline{E}_k} = \delta^{\underline{P}}_{\underline{E}_k}, \qquad \delta^{\underline{E}_k}_{\underline{R}} = \delta^{\underline{E}_k}_{\underline{P}},$$

for all $k \in \{1, \dots, N\}$. For example, take $N = 1$ and any $\underline{P} \not\leq \underline{P}_1$. For every reasonable choice of $\underline{R}$, that is, $\underline{R} = \underline{P}_1$ or $\underline{R} = \inf_{\omega\in\Omega}$, it *cannot* hold that

$$\underline{P}_1 \geq \underline{R} \implies \underline{P}_1 \geq \underline{P}.$$

There does not seem to exist an efficient method for calculating neither $\underline{E}^u$, $\bar{\mathfrak{t}}^e(\underline{Q})$, nor $\underline{\mathfrak{v}}^e(\underline{Q})$.

6.1 An algorithm for calculating the first-order aggregate

In the following we assume that $\bar{\mathfrak{v}}$ is 1, and $\underline{R}_j$, $j = 0, \dots, M$ denotes an enumeration of all possible $\sqcap_{k\in K}\underline{P}_k$, $K \subseteq \{1, \dots, N\}$. Without loss of generality we may assume that $\underline{R}_0 = \inf_{\omega\in\Omega}$. Theorem 1 shows that, in some cases the natural extension can be calculated by solving a finite linear program (an infinite number of linear inequalities reduces to a finite number of linear inequalities). In its dual form, this linear program has a very nice form, it is given in the next theorem.

Theorem 2 (Dual form). *Let $X \in \mathcal{L}(\Omega)$ be any (first-order) gamble. Assume that $\bar{\mathfrak{v}}$ is 1, and let $\underline{R}_j$, $j = 0, \dots, M$ denote an enumeration of all possible $\sqcap_{k\in K}\underline{P}_k$, $K \subseteq \{1, \dots, N\}$. Without loss of generality we assume that $\underline{R}_0 = \inf_{\omega\in\Omega}$. Let $\alpha_0, \dots, \alpha_M$ be non-negative variables that maximize $\sum_{j=0}^M \alpha_j \underline{R}_j(X)$ subject to*

$$\sum_{j=0}^M \alpha_j = 1 \qquad \sum_{j=0}^M \alpha_j \delta_{\underline{E}_j}^{\underline{E}_k} \geq \underline{\mathfrak{v}}(\underline{E}_k)$$
$$\sum_{j=0}^M \alpha_j \delta_{\underline{E}_k}^{\underline{R}_j} \geq \underline{\mathfrak{t}}(\underline{E}_k) \qquad \sum_{j=0}^M \alpha_j \delta_{\underline{E}_k}^{\underline{R}_j} \leq \bar{\mathfrak{t}}(\underline{E}_k)$$

for all $k \in \{1, \dots, M\}$. (Notice that the constraints do not depend on the gamble X). Then

$$\underline{E}^\ell(X) = \sum_{j=0}^M \alpha_j \underline{R}_j(X).$$

6.2 Practical setup

We have described one way to formulate and solve the algebraic problem set. Practically, each problem can be solved using the following strategy. Each expert k specifies a lower prevision $\underline{P}_k$ on a finite subset $\mathcal{K}_k$ of gambles. The modeller associates lower trust t_i^ℓ, upper trust t_i^u, and lower tsurt v_i^ℓ with each expert k. Use Proposition 1 to enumerate all $\underline{R}_k(X) = \sqcap_{k\in K}\underline{P}_k(X)$ for any gamble X of interest. Also determine $\delta_{\underline{E}_k}^{\underline{R}_j}$ and $\delta_{\underline{R}_j}^{\underline{E}_k}$ for every $k \in \{1, \dots, N\}$ (point-wise dominance can also be tested efficiently by means of a linear program). Finally, use Theorem 2 to calculate $\underline{E}^\ell(X) = \underline{E}(X_*)$ for any gamble X of interest.

6.3 Example: Poincaré's paradox

Poincaré's paradox [12] arises when we consider three objects a, b and c, such that a cannot be distinguished from b, b cannot be distinguished from c, but

Table 1: $\delta_{\underline{P}_i}(\underline{R}_j)$ for Poincaré's Paradox

	$a = b$	$b = c$	$a \neq c$
no assessment	0	0	0
$a = b$	1	0	0
$b = c$	0	1	0
$a \neq c$	0	0	1
$a = b \wedge a \neq c$	1	0	1
$b = c \wedge a \neq c$	0	1	1
$a = b \wedge b = c$	1	1	0

clearly a is not equal to c. It thus consists of the assessments $a = b$, $b = c$ and $a \neq c$. We investigate to what extent these assessments are consistent within the present approach.

To this end, we assign a lower trust t to each assessment. The conjunctions $\underline{R}_j$ and the coefficients $\delta_{\underline{P}_i}(\underline{R}_j)$ are listed in Table 1. The corresponding linear programming problem has a feasible solution only for $t \leq \frac{2}{3}$. This means that Poincaré's paradox can be resolved only if we trust each assessment up to a degree of $66,7\%$. For $t = \frac{2}{3}$ the solution is $\underline{E}^{\ell} = \frac{1}{3}(\underline{P}_{a=b\wedge a\neq c} + \underline{P}_{b=c\wedge a\neq c} + \underline{P}_{a=b\wedge b=c})$. Since this conclusion only depends on the values of the $\delta_{\underline{P}_i}(\underline{R}_j)$, any three conflicting assessments that are pair-wise consistent, are actually consistent up to an equally distributed degree of trust of $66,7\%$.

7 Discussion and conclusion

A second-order imprecise probability model was proposed based on a behavioural notion of trust. As most second-order hierarchical models, the interpretation of this model relies on the existence of a hypothetical "representative" expert. Unluckily, this leads to philosophical as well as to practical problems. The second-order gambles that were used to derive the aggregate are defined on a possibility space that cannot always be sampled in a meaningful way. How should one deduce the second-order lower and upper trust and tsurt values? One could argue that the model should only be used in applications where the representative expert can be identified (for instance, one choice could be identifying the representative expert with the modeller itself).

Despite these problems, I think the method is both mathematically simple, sufficiently general, and practically appealing, especially if only a limited number of expert opinions need to be aggregated. Theorem 2 shows that the first-order aggregate can be obtained as a weighting rule: it is a convex combination of conjunctions of all possible (non-conflicting) combinations of experts. The weights may still depend on the gamble of interest, however. In the general case however, the size of the linear program to be solved will grow exponentially in the number of experts, limiting the applicability of the model. Also, there seems to be no efficient method for calculating neither $\underline{E}^u$, $\overline{\mathfrak{t}}^e(\underline{Q})$, nor $\underline{\mathfrak{v}}^e(\underline{Q})$.

Restricting to lower trust only, it is easy to obtain the following results (the lower trust assignments are assumed to be non-negative):

(i) If lower trust is equal to one for all experts, then the first order aggregate is equal to the conjunction of all the experts, and second-order consistency is equivalent with consistency.

(ii) If a lower trust model is second-order consistent, then it will remain so for any lower assignment of lower trust.

(iii) A lower trust assignment such that $\sum_{k=1}^{N} t_k^\ell \leq 1$ is always second-order consistent: in that case, a first order aggregate always exists.

(iv) If all experts are pair-wise conflicting, that is, if there are no conjunctions except for the trivial ones, then any lower trust assignment such that $\sum_{k=1}^{N} t_k^\ell > 1$ is second-order conflicting. If $\sum_{k=1}^{N} t_k^\ell \leq 1$, then there is second-order consistency, and the aggregate is given by

$$\underline{E}^\ell(X) = \sum_{k=1}^{N} t_k^\ell \underline{E}_k(X) + \Big(1 - \sum_{k=1}^{N} t_k^\ell\Big) \inf_{\omega\in\Omega} X(\omega).$$

Thus in case of total conflict (which is quite common if all experts use a single probability measure to represent their knowledge), the model produces a generalised linear opinion pool.

These results show that the highest possible assignments for lower trust measure the amount of conflict between the expert assessments. If they can be chosen maximal, all equal to one, then no conflict is present. If they cannot even be chosen such that their sum is larger than one, then there is total conflict.

Acknowledgments

The author would like to thank Gert de Cooman for his stimulating discussions and comments.

References

[1] Jean-Marc Bernard, Teddy Seidenfeld, and Marco Zaffalon, editors. *ISIPTA '03 – Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications.* Carleton Scientific, 2003.

[2] G. de Cooman, F. G. Cozman, S. Moral, and P. Walley, editors. *ISIPTA '99 – Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*, Ghent, 1999. Imprecise Probabilities Project.

[3] Gert de Cooman. Lower desirability functions: A convenient imprecise hierarchical uncertainty model. In de Cooman et al. [2], pages 111–120.

[4] Gert de Cooman. Precision-imprecision equivalence in a broad class of imprecise hierarchical models. *Journal of Statistical Planning and Inference*, pages 175–198, June 2002.

[5] Gert de Cooman and Matthias C. M. Troffaes. Coherent lower previsions in systems modelling: products and aggregation rules. Accepted for publication in Reliability Engineering and System Safety, 2003.

[6] Gert de Cooman and Peter Walley. A possibilistic hierarchical model for behaviour under uncertainty. *Theory and Decision*, 52(4):327–374, 2002.

[7] Bruno de Finetti. Sul significato soggettivo della probabilità. *Fundamenta Mathematicae*, 17:298–329, 1931.

[8] Bruno de Finetti. *Theory of Probability: A critical introductory treatment*. Wiley, London, 1975.

[9] Christian Genest and James V. Zidek. Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114–148, 1986.

[10] S. Moral and J. del Sagrado. Aggregation of imprecise probabilities. In B. Bouchon-Meunier, editor, *Aggregation and Fusion of Imperfect Information*, pages 162–188. Physica-Verlag, New York, 1998.

[11] Robert F. Nau. The aggregation of imprecise probabilities. *Journal of Statistical Planning and Inference*, 105(1):265–282, June 2002.

[12] Henri Poincaré. *La science et l'hypothèse*. Flammarion, Paris, 1902.

[13] Lev V. Utkin. A second-order uncertainty model of independent random variables: An example of the stress-strength reliability. In Bernard et al. [1], pages 532–546.

[14] Lev V. Utkin and Thomas Augustin. Decision making with imprecise second-order probabilities. In Bernard et al. [1], pages 547–561.

[15] P. Walley. The elicitation and aggregation of beliefs. Technical report, University of Warwick, Coventry, 1982. Statistics Research Report 23.

[16] P. Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83:1–58, 1996.

[17] P. Walley. Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems*, 26:337–383, 1997.

[18] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[19] N. Wilson and S. Moral. Fast markov chain algorithms for calculating dempster-shafer belief. In W. Wahlster, editor, *Proceedings of the 12th European Conference on Artificial Intelligence*, pages 672–676. J. Wiley, 1996.